



Report on planned modifications

Team MaLi

Overview - Problems

- Syntax/word order

1. ID: 161

2. *Lühike ent vägivaldne Araabia-lisraeli sõda 1973. aasta oktoobris toob Euroopas kaasa energiakriisi ning majandusprobleemid.*

3. *The short, yet brutal, Arab-Israeli war of October 1973 result in an energy crisis and economic problems in Europe.*

4. *A short but violent Arab war in October 1973 brings the energy crisis into Europe and economic problems.*

- Incorrect grammar tense

1. ID: 438

2. *Neil oli palju probleeme raadioga võistlemisel, kuid televisioon oli jällegi hoopis teisest klassist.*

3. *They had had many problems in competing with radio, but television was in a different league again.*

4. *They had a lot of trouble competing with the radio, but the TV was another class.*

- Wrong translations regarding the sense

1. ID: 951

2. *1. Klõpsake kommentaari jutumulli, mida soovite redigeerida.*

3. *1. Click inside the balloon for the comment that you want to edit.*

4. **One click on the legend of the comment you wish to make**

Overview - Problems

- Missing words

1. ID: 854

2. *Luuu integreeritud, paindlikku teenusepõhist arhitektuuri*

3. *Create an integrated, flexible service-oriented architecture*

4. *Create an integrated, flexible architecture architecture*

- Word-by-word translation, like "black mass" instead of "dark matter"

1. ID: 801

2. *Nad seetõttu postuleerivad nähtamatu musta massi olemasolu, mis on vastastikmõjus muu mateeriaga ainult selle gravitatsiooni kaudu, ja sisaldab väga stabiilseid, seega ka väga kergeid osakesi.*

3. *They therefore postulate the existence of an invisible dark matter that only interacts with the rest of matter by its gravitation, and consists of very stable, and therefore very light, particles.*

4. *They are therefore exposed to the presence of an invisible black mass that is in interaction with other matter only through its gravity, and contains very stable, so very mild particles.*



Solutions

Back translation

1. ID: 161

2. Lühike ent vägivaldne Araabia-lisraeli sõda 1973. aasta oktoobris toob Euroopas kaasa energiakriisi ning majandusprobleemid.

3. The short, yet brutal, Arab-Israeli war of October 1973 result in an energy crisis and economic problems in Europe.

4. A short but violent Arab war in October 1973 brings the energy crisis into Europe and economic problems.

- Improve quality/fluency of translation
- Automatic translation of the monolingual target text → additional parallel training data
- No changes to neural network architecture → easy to apply

Replace BPE with “wordpieces”

- several problems regarding BPE splitting

1. ID: 530

2. *Ungaris leiti, et peaaegu 96% lampidest on ohtlikud.*

3. *In Hungary, nearly 96% of the lights were found to be hazardous.*

4. *In Hungary, almost 96% of sheep were found to be dangerous.*

- → „wordpiece”
 - Variant of BPE
 - Final vocabulary size is fixed

Hyperparameters – Beam Search

1. ID: 182

2. Riikidevaheline reisimine on lihtsam, kuna enam pole rahavahetuse kulu ja vaeva.

3. Travel between euro countries is easier because the costs and inconvenience of changing money have disappeared.

4. Transnational travel is easier, since there is no more money for money and work.

- Could help to choose more appropriate words
- Change beam from default 5 to 10
- Involves candidates which are still close to the „best one“

Transformer Model

3. *She urged EU member countries to do a better job of enforcing safety standards and said industry – manufacturers, retailers, distributors and importers – should also step up their checks.*
4. *He urged EU Member States to do better work on security requirements and said that industry (producers, retailers, dealers and importers) should also make controls more effective.*

- Replaces recurrent layers with self-attention
- Compares each word to every other word in the sentences
- Dependencies between distant positions
- Requires less time to train